



## Review

# An unsolved mystery: The target-recognizing RNA species of microRNA genes


Chang-Zheng Chen <sup>a,b,\*</sup>
<sup>a</sup> Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA

<sup>b</sup> Baxter Laboratory for Stem Cell Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

## ARTICLE INFO

## Article history:

Received 1 February 2013

Accepted 6 May 2013

Available online 14 May 2013

## Keywords:

MicroRNA

Antisense regulatory RNA microRNA

Primary microRNA

Precursor microRNA

Mature microRNA

Target recognition

## ABSTRACT

MicroRNAs (miRNAs) are an abundant class of endogenous ~21-nucleotide (nt) RNAs. These small RNAs are produced from long primary miRNA transcripts – pri-miRNAs – through sequential endonucleolytic maturation steps that yield precursor miRNA (pre-miRNA) intermediates and then the mature miRNAs. The mature miRNAs are loaded into the RNA-induced silencing complexes (RISC), and guide RISC to target mRNAs for cleavage and/or translational repression. This paradigm, which represents one of major discoveries of modern molecular biology, is built on the assumption that mature miRNAs are the only species produced from miRNA genes that recognize targets. This assumption has guided the miRNA field for more than a decade and has led to our current understanding of the mechanisms of target recognition and repression by miRNAs. Although progress has been made, fundamental questions remain unanswered with regard to the principles of target recognition and mechanisms of repression. Here I raise questions about the assumption that mature miRNAs are the only target-recognizing species produced from miRNA genes and discuss the consequences of working under an incomplete or incorrect assumption. Moreover, I present evolution-based and experimental evidence that support the roles of pri-/pre-miRNAs in target recognition and repression. Finally, I propose a conceptual framework that integrates the functions of pri-/pre-miRNAs and mature miRNAs in target recognition and repression. The integrated framework opens experimental enquiry and permits interpretation of fundamental problems that have so far been precluded.

© 2013 The Author. Published by Elsevier Masson SAS. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. An unsolved puzzle: the target-recognizing RNA species from miRNA genes

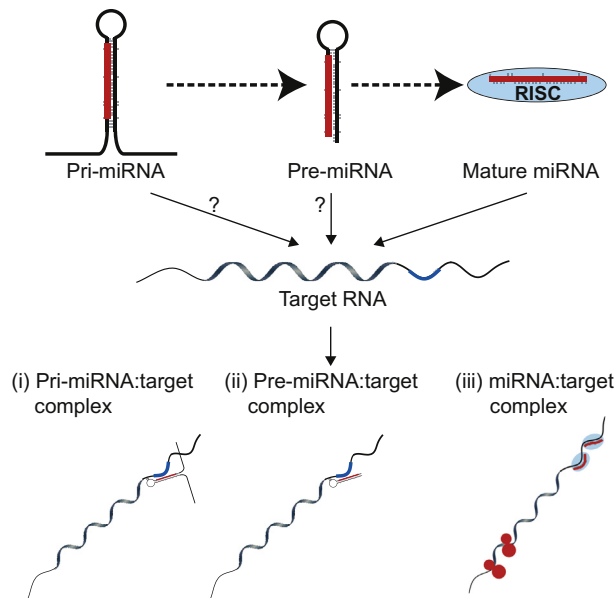
The assumption underlying all but a few miRNA studies to date [1–3] is that mature miRNAs are the only target-recognizing RNA species produced from miRNA genes [4–6]. This “first principle” of the miRNA field has guided research design and data interpretation of the field for more than a decade (for reviews, see Refs. [7–9]). However, Nature presented us with an intriguing and deceiving puzzle: The major RNA species made from a miRNA gene – the pri-, pre-, and mature miRNAs – each contain a region of identical sequence (that of the mature miRNA), and, therefore, all have the

potential to interact with target mRNAs (Fig. 1). As all three forms may be present in the same cells, it is intrinsically difficult to distinguish their physiological contributions to target recognition and repression by either biochemical or genetic approaches. No experiments have proven that mature miRNAs are the only target-recognizing RNA species produced from miRNA genes and no studies have definitively ruled out functions of pri-/pre-miRNAs in target recognition and repression.

The above discussion immediately raises questions as to how the first principle of miRNA field was established initially. To answer these questions, one has to thoroughly review the emergence and evolution of the miRNA field and that of the parallel field RNA interference (RNAi) [4–6,10]. In this article I will review the key evidence that, in my opinion, contributed to acceptance of mature miRNAs as the sole functional RNA products of miRNA genes. Moreover, I will discuss why existing studies fall short of establishing that mature miRNA are the only target-recognizing miRNA species and why it is essential to determine the roles of pri-, pre-, and mature miRNA species in target recognition and repression.

\* Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. Tel.: +1 650 725 1667.

E-mail address: [czchen@stanford.edu](mailto:czchen@stanford.edu).



**Fig. 1.** The puzzle: the target-recognizing species produced from miRNA genes? The major RNA products of miRNA genes – pri-miRNA, pre-miRNA, and mature miRNA – all contain the ~21 nucleotides found in the mature miRNA, and, in principle, all can interact with the target mRNAs. Genetic and biochemical analyses carried out so far do not establish mature miRNAs as the only target-recognizing molecules made from a miRNA gene or rule out a function for pri-/pre-miRNAs in target recognition. Thus, it is still unknown to what extent pri-/pre-/mature miRNAs contribute the activity of miRNA genes *in vivo*. Ignoring this puzzle will misdirect research as: (1) only partial information will be used for target prediction; (2) targets will be sought with incorrect RNA species as bait; (3) regulatory proteins will be assigned to incorrect pathways; and (4) experimental design and interpretation will be biased.

### 1.1. The importance of defining all target-recognizing miRNA species

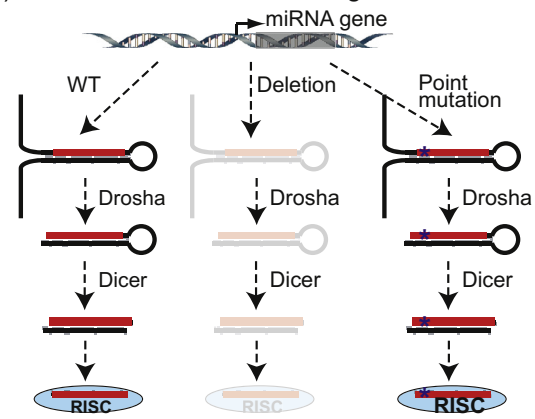
Resolving the functions of miRNAs and their precursor species in target recognition is a prerequisite for elucidating the mechanisms and principles underlying target recognition and repression. Without this knowledge, it is not possible to decipher the precise information encoded in the miRNA genes that is utilized in target recognition (Fig. 1). For example, if mature miRNA is the sole target-recognizing species, then only the ~22 nucleotide linear sequence of the mature miRNAs are utilized for target recognition. Thus, for target identification one would develop experimental approaches to identify mRNAs that interact with mature miRNAs, and for target prediction one would only use the nucleotide sequence within mature miRNAs for computational analyses. In contrast, if pri- and/or pre-miRNAs play roles in target recognition and repression, these miRNA precursors should be used for experimental and computational target identification. Similarly, without definitive knowledge about the target-recognizing miRNA species, it is not possible to decipher the molecular and cellular pathways that are integral to target recognition and repression. In particular, it would cause confusion or controversies to try to fit the molecular pathways that are important for pri- and/or pre-miRNA functions onto the mature miRNAs [11–14]. Given the consequences, it is important to examine the evidence and arguments that led to prevailing mature miRNA model.

### 1.2. Genetic analyses cannot discern the target-recognizing miRNA species

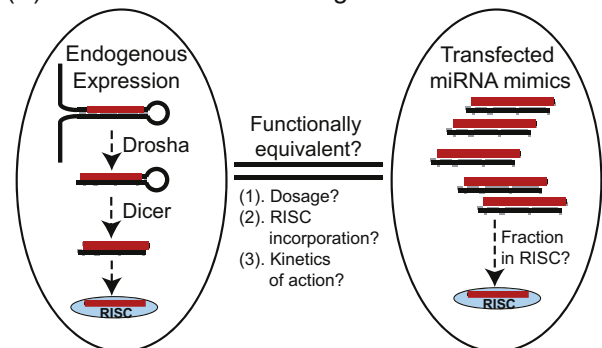
Genetic analyses, which led to the discovery of miRNA genes and subsequent functional characterization in animals, are not

sufficient to discern the functional contributions by mature and pri- and/or pre-miRNAs in target recognition (Fig. 2A). Loss-of-function analyses (i.e., deletion or mutation of miRNA genes) results in simultaneous disappearance or alteration of all three miRNA species – pri-, pre-, and mature miRNAs. Therefore, it is not possible to attribute the phenotypes of such loss-of-function analyses to one of the RNA species produced by a miRNA gene. Similarly, gain-of-function analyses – as performed through expression of genes encoding primary miRNAs – suffer the same limitations since all three RNA species are produced. This argument was succinctly presented in the original discoveries of the first miRNA gene – *lin-4* [15]. At that time, the primary *lin-4* transcripts had not been

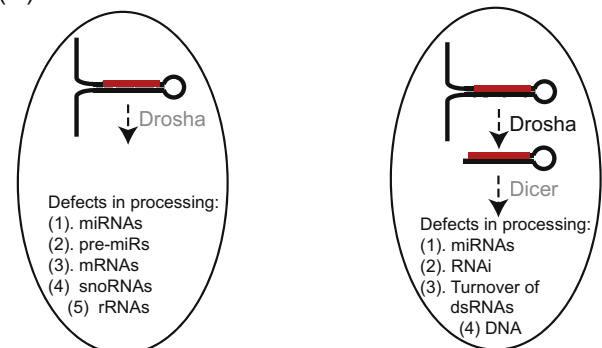
### (A). Genetic ablation of miRNA genes



### (B). miRNA mimics as surrogates:



### (C). Drosha or Dicer null cells



**Fig. 2.** Experiments that led to the canonical miRNA models and their limitations. (A) Limitations of genetic deletion and mutation analyses for determination of the target-recognizing RNAs made from a miRNA gene. (B) Transfected mature miRNA mimics are not functionally equivalent to the endogenous miRNA gene products. (C) Both *drosha*- and *dicer*-null cells have broad defects beyond those resulting from lack of production of a single mature miRNA.

detected, and it was postulated that either precursor or mature *lin-4* miRNAs (referred to as *lin-4L* and *lin-4S*, respectively, in the original paper) or both could be the functional *lin-4* gene products. The authors favored the hypothesis that the mature *lin-4* miRNA (*lin-4S*) played the major role in base pairing with *lin-14* target RNA, since mature *lin-4* miRNA is significantly more abundant than precursor *lin-4* miRNA (*lin-4L*) and because the secondary structure of *lin-4L* might sequester bases required for interactions with the target mRNA. Genetic studies of miRNA gene function in worms, zebrafish, and mammals have the same limitations (Fig. 2A). Therefore, although genetic analyses have revealed diverse effects of gain and loss of miRNA genes in normal and pathogenic animal physiology, such analyses cannot discern the target-recognizing miRNA species or be used to attribute these phenotypes to the mature miRNAs only.

### 1.3. miRNA mimics cannot be used to preclude pri-/pre-miRNA function in target recognition

The seminal discoveries in the fields of RNA interference [16] and miRNAs might have propelled the convergences of the two fields. RNAi was shown to be mediated by small interfering RNAs (siRNAs) of ~25 nt in plants [17] and ~21 nt in animals [10,18,19], which are remarkably similar to miRNAs [4–6,10,20]. Both siRNAs and miRNAs are the products of a conserved ribonuclease III enzyme Dicer [21,22] and are integral components of the RISC [23,24]. Synthetic siRNA duplexes have been shown to function as miRNAs *in vitro* and *in vivo* [20,25]. Thus, both siRNAs and miRNAs function to guide RISCs to cognate mRNA targets and control their expression, supporting that hypothesis that mature miRNAs are a functional RNA species. However, it is not known whether transfected mature miRNAs functionally mimic the endogenous miRNA gene products in terms of degree of activity, mode of action, and range of targets (Fig. 2B). Moreover, it is difficult to compare the *in vivo* biological activities of endogenous and transfected mature miRNAs since effects of transfection are transient. Finally, most importantly, transfection of miRNA mimics cannot rule out the function of miRNA precursor species. Thus, miRNA mimics cannot be used to establish that mature miRNAs are the only target-recognizing miRNA genes or to rule out the functions of pri-/pre-miRNAs in target recognition.

### 1.4. The limitations of dicer- or drosha-null cells in discerning the physiological functions of precursor and mature miRNA species

Mature miRNAs are generated through stepwise processing of miRNA precursors by the ribonuclease III enzymes Drosha in the nucleus and Dicer in the cytosol [8]. Deletion of Drosha or Dicer in cells blocks the processing of pri-miRNA and pre-miRNAs, respectively, and causes defects in mature miRNA biogenesis. However, Dicer and Drosha are known to have roles beyond those in mature miRNA biogenesis. The Drosha complex binds and processes mRNAs, snoRNAs, and lincRNAs [27,28]. Dicer-1 plays an essential role in RNA interference by generating siRNAs from foreign and endogenous sources of double-stranded RNAs [21] and can also fragment chromosomal DNA during apoptosis [29]. Furthermore, these enzymes are responsible for the generation of hundreds of mature miRNAs. Therefore, deletion of either Dicer or Drosha results in severe and pleiotropic defects that cannot be attributed to the loss of a single mature miRNA. A number of studies have demonstrated that transfected miRNA mimics may partially rescue the defects caused by disruption miRNA biogenesis machinery, including inhibition of Dicer or Drosha/GDCR8 expression [25,26]. Although these experiments indicate that transfected mature miRNAs have some functional roles in the processes examined

(Fig. 2B), they do not demonstrate that transfected miRNA mimics are functionally equivalent to the endogenous mature miRNAs or rule out the function of pri- and/or pre-miRNAs.

Moreover, the *dicer* or *drosha*-null cells are likely to have limited uses in discerning the physiologic functions of pri-miRNAs or pre-miRNAs. First, it is not known whether deletion of Drosha or Dicer cause cellular defects that compromise the functions of pri-/pre-miRNAs. Second, if pri- or pre-miRNAs have functions independent of mature miRNAs, disruption of mature miRNA biogenesis through deletion of *dicer* or *drosha* genes may affect the stability and the concentration of pri- or pre-miRNAs and influence their activities. Finally, pri- or pre-miRNA-mediated gene regulation may be a multistep process and directly or indirectly require Dicer and/or Drosha enzymes or their associated complexes (see Section 4, [1–3]). Given the limitations, it is important to interpret the observations made using the *dicer*- or *drosha*-null cells with care.

In summary, the discovery of *lin-4*, *let-7*, and many other miRNAs helped to establish a general case for the presence of many similar noncoding RNA genes in animal kingdoms that produce small ~22-nt miRNAs. However, the canonical miRNA model, which was built on the original discovery of *lin-4* and *let-7* miRNA genes in *Caenorhabditis elegans* [15,30], has clear limitations. An inclusive framework is necessary to resolve the functional contributions of the major RNA species produced from miRNA genes and to develop better computational and experimental approaches for target identification.

## 2. Lessons from bacterial antisense regulatory RNAs

Should we even consider that pri-/pre-miRNAs function in target recognition and repression? Could the secondary structure of pri-/pre-miRNAs sequester most of the complementary bases complementary, rendering them inactive for basing pairing with target mRNAs [15]? In this section, I will use bacteria antisense regulatory and other examples from the RNA World to illustrate how regulatory information encoded in RNA structures can be directly translated into activities through interactions between structured RNA molecules.

### 2.1. RNAs interact through structured motifs

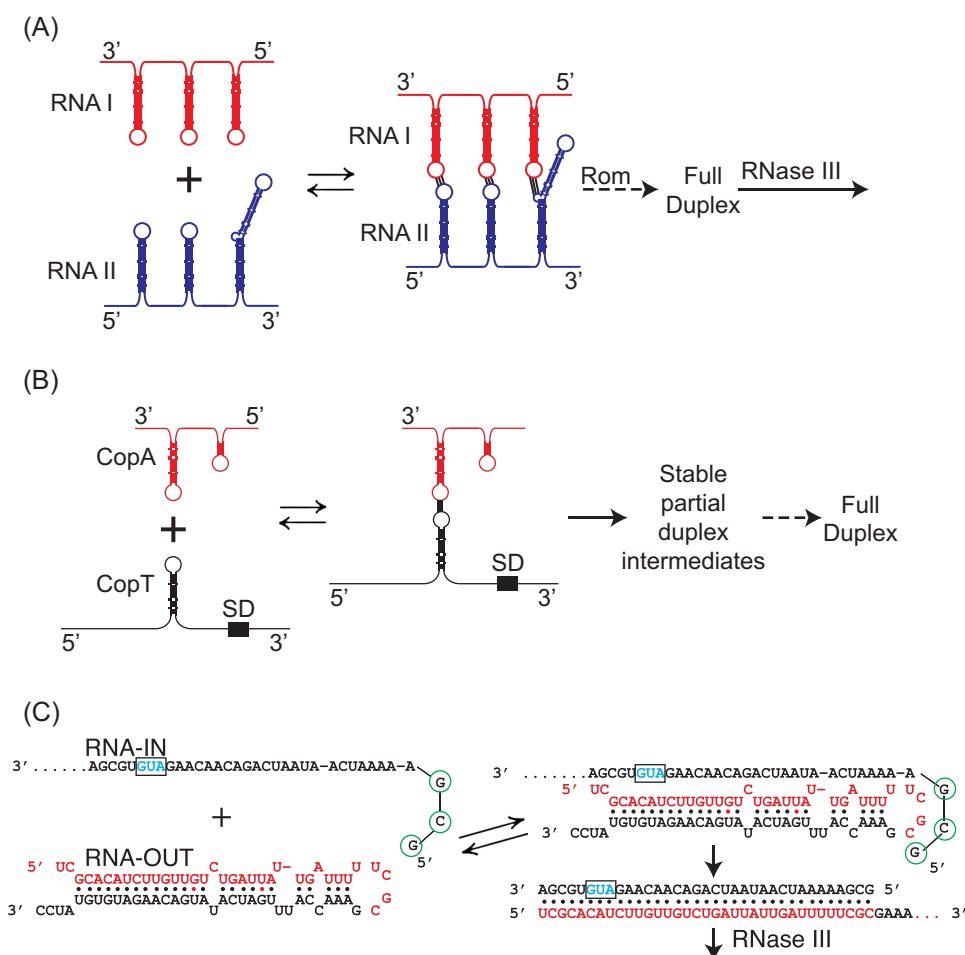
Notably, most RNA molecules do not exist in linear forms. They fold into secondary and tertiary structures under physiological conditions independent of proteins or as integral components of ribonucleoprotein (RNP) complexes (for reviews, see Refs. [31–34]). The structural constraints on RNA sequence elements within these folded RNAs or RNPs are essential for their activities in catalysis and molecular recognition, whereas the same linear RNA sequence elements without structural constraints or with incorrect constraints are generally ineffective. For example, complexes formed between two tRNA anticodon loops are 100–10,000-fold more stable than those formed between a linear RNA and tRNA loop [35]. The stacking of bases surrounding the recognition sequences in these loops and the constraints on the loop structures contribute to the higher stability of complexes formed between structured RNAs than linear strands. It is important to note that structure-based RNA:RNA interactions are not only critical for tRNA anticodon interactions with mRNA [36], but are also important for plasmid replication [37], transcription anti-termination [38], dimerization of retroviral genomic RNA [39], and recognition of splice junctions by small nuclear RNAs in the spliceosome [40] to name only a few of the biological functions mediated by structured RNAs. In short, regulatory RNAs generally utilize structure-based interactions to recognize their targets and to control diverse molecular processes such as replication, transcription, translation, splicing, and DNA recombination. Various functions and

mechanisms of antisense RNA regulation have been reviewed elsewhere [31–34,41]. I will discuss a few well-established examples of bacterial antisense RNAs to illustrate some of the general properties of structured RNAs in gene regulation and their similarities to the structured pri- and pre-miRNAs.

## 2.2. Control of plasmid copy numbers by antisense regulatory RNAs

Bacteria antisense RNAs provide the prototypic illustration of gene regulation by structured antisense RNAs. The first example of antisense regulation was discovered during the study of replication control of plasmid ColE1 in *Escherichia coli* [31]. ColE1 and related plasmids encode two noncoding RNAs, RNA I and RNA II. Transcription and processing of RNA II generates the primers required for the initiation of plasmid replication. RNA I, a partially overlapping antisense transcript of RNA II, negatively controls the production of replication primers through direct interaction with RNA II. Higher plasmid concentrations result in increased RNA I synthesis which then prevents primer formation by RNA II and inhibits replication. Both RNA I and RNA II form stem-loop structures with seven nucleotides in the loops and 5–20 base pairs in

the stems (Fig. 3A). The base pairing between RNA I and II is initiated by a transient and reversible interaction, dubbed a “kissing” interaction, between complementary nucleotides of the three stem-loops of RNA I and RNA II [42]. The kissing of the loop regions brings the two antisense RNAs together and promotes the formation of thermodynamically stable complexes by propagating base pairing and formation of full duplexes. Biochemical evidence indicates that the rate-limiting step in forming the stable complexes is the kissing step [43]. Both nucleotide sequences within the loops and their structural constraints impact the association rates and the stabilities of resultant complexes. Notably, the RNA I modulator Rom, a 63 amino acid protein, binds to the RNA I/RNA II complex and decreases the dissociation rate by over 100-fold [44]. The efficiency of ColE1 replication correlates well with the association rates of the complexes as well as rates of synthesis and the concentrations of RNA I, RNA II, and Rom [45]. Plasmid copy number has an inverse correlation with the association rates of RNA I and II. Nucleotide changes in the stem-loops of RNA I also affect plasmid concentration and incompatibility [46,47], both the activity and specificity of the RNA I are determined by the loop sequences and the stem-loop structures of the RNA.



**Fig. 3.** Bacteria antisense regulatory RNAs. (A) Both RNA I and RNA II form stem-loop structures and their interaction is controlled by formation of “kissing” intermediates. Formation of the kissing complexes induces more extended base pairing between the two perfectly matched antisense RNAs through a multi-step process that results in long duplexes. RNA I, RNA II, and the fully formed duplexes are substrates of RNase III, which has an evolutionarily conserved role in processing double-stranded RNAs. Rom potentiates the formation of RNA I and RNA II complexes. (B) Control of R1 plasmid copy number of the IncFII family is in part mediated through the antisense regulation of the synthesis of RepA protein, which is required for R1 replication. Binding of antisense RNA CopA with its target CopT (RepA mRNA), which is controlled by kissing complex formation, results in the formation of double-stranded RNA and blocks RepA translation. CopA, CopT, and their duplexes are sensitive to RNase III. (C) The frequency of IS10 transposition is controlled by the expression of IS10-encoded transposase (RNA-IN), which is in part regulated by the antisense RNA-OUT. RNA-OUT forms a stem-loop structure and interacts with the linear region at the 5' end of RNA-IN. The initial recognition of RNA-IN by RNA-OUT is mediated by three nucleotides of RNA-OUT (highlighted in green circles) and leads to the formation of a fully paired duplexed complex that are degraded by RNase III.



Diverse mechanisms underlie the regulatory effects of antisense RNAs in bacteria. Binding of antisense RNAs can also inhibit translation by inducing premature transcription termination of the target mRNA, rendering the target susceptible to digestion by RNase III or other ribonucleases, or by blocking the ribosomal binding site and/or the initiation codon (see reviews [31–33,48]). For example, the synthesis of the RepA protein from the R1 plasmid is inhibited by the binding of antisense RNA CopA with its target CopT [41,49,50], resulting in the formation of double-stranded RNA that is sensitive to RNase III cleavage (Fig. 3B). Moreover, antisense inhibition of translation plays critical roles in controlling the transposition of Insertion Sequence 10 (IS10) [32]. Specifically, IS10 encodes antisense RNA-OUT that is complementary to the 35 nucleotides at the 5' end of the transposase mRNA, RNA-IN. The interaction of RNA-OUT with RNA-IN blocks access to the start codon and ribosome-binding site and inhibits the translation of transposase (Fig. 3C). In both cases, complex formation between the antisense RNAs and their regulatory targets is initiated by the kissing between the antisense stem–loop and the complementary structured RNA target (i.e. CopT) or linear RNA target (RNA-IN). Again, both the nucleotides within these loops and their structural constraints are critical for this antisense-mediated regulation and the antisense RNA and target complexes are substrates of RNase III enzymes *in vivo*.

Finally, chromosomally encoded antisense RNAs, which have partial complementarity to their target RNAs, control gene expression by the formation of short and/or imperfect duplexes [51]. Among these is an oxidative stress induced non-coding RNA, OxyS, which helps protect bacterial cells against DNA damage by targeting *fhfA* mRNA through a formation of a kissing complex intermediate facilitated by the Hfq protein [3,52–55]. Generally, the interactions of target mRNAs and bacterial regulatory RNAs, whether they have perfect or imperfect antisense matches to their cognate target mRNAs, are mediated by stem–loop motifs and formation of kissing complexes [34].

### 2.3. Insights from bacteria antisense regulatory RNAs

These examples illustrate that antisense regulatory RNAs are utilized to measure and adjust the cellular concentrations of DNA, RNA, and protein. The modes of action by these naturally occurring antisense RNAs illustrate the evolutionarily conserved principles of gene regulation that may be applicable to regulatory RNAs in other branches of life. The principles underlying effective antisense regulation learned from the bacteria antisense RNAs were thoroughly described in excellent review articles by Wagner et al. [56] and Brunel et al. [33]. Here I emphasize on one of the fascinating features of bacterial antisense regulatory RNAs; that is, they invariably fold into structures with characteristic stem–loops that are critical for their function. Moreover, despite the extensive base pairing potential between the antisense RNAs and their targets, all characterized examples indicate that a few structurally constrained nucleotides within stem–loops dictate the initial interactions between antisense regulatory RNAs and their targets. These examples illustrate an ancient mechanism through which the regulatory information encoded in structured antisense RNAs can be directly translated into function through interactions with their regulatory targets. These examples from the RNA World lend support to the notion that structured pri-/pre-miRNAs interact with target mRNAs.

Why has evolution selected structured RNA motifs rather than the linear antisense nucleotides to control the interaction between antisense regulatory RNAs and their cognate targets? What are the advantages of structure-based RNA interactions over the interactions completely based on linear sequences? One possible

answer is that structured antisense RNAs provide mechanisms that control the specificity of interactions that are not available through the simple linear RNA sequence [31,57]. First, structure apparently prevents non-specific linear interactions between antisense RNAs. Direct interaction between long linear antisense RNAs has the tendency to form strong but non-specific interactions since relative short RNA duplexes can be very stable and tolerate bulges and gaps. These non-specific complexes, once formed, are difficult to readjust. Folding of RNA into thermodynamically stable secondary and tertiary structures can prevent such non-specific linear interactions. Moreover, the formation of initially transient and unstable kissing complexes between the antisense RNAs serves as a scanning mechanism to facilitate the antisense regulatory RNAs to find the correct binding sites and to promote the formation of stable and extended duplexes. Importantly, minor changes in loop nucleotides or the structural constraints (e.g., changes in the identity of loop nucleotides or alteration of the stacking of loop nucleotides by changes in loop size or stem base pairs) influence the specificities and rates of kissing complex formation and have drastic effects on target recognition by the structured antisense RNAs. Thus, the kissing mechanism provides unprecedented specificity control in target recognition and allows target specificities to evolve through modification of stem–loop sequence and structure.

Finally, cellular proteins, like ribonucleases such as RNase III and Hfq, may recognize antisense RNA, target RNA, and their complexes, thereby controlling their stability and the activity of antisense RNAs. For example, Spot 42, an Hfq-binding bacteria small RNA, can regulate large number of targets through coordinated recognition by the antisense RNA and Hfq protein [58]. Target structure can affect antisense and/or Hfq recognition of the cognate targets, thereby influencing outcomes of the regulation. In summary, as illustrated by target recognition by the RNA I, CopA, RNA-IN, tRNA, and examples discussed elsewhere, structured RNAs offer precise control of the sensitivity, strength, and specificity of RNA:RNA interactions that are not available from simple linear RNA interactions. Thus, the intrinsic properties of a large RNA molecule to form secondary and tertiary structure provide many advantages for target recognition by antisense RNAs [31–34].

## 3. An integrated framework of regulatory controls by miRNA gene products

The above discussion illustrates that it is the intrinsic properties of RNA molecules to encode regulatory information in their secondary and tertiary structures, and such information can be directly interpreted in gene regulation through RNA:RNA interaction as exemplified by the bacteria antisense regulatory RNAs. Intuitively, structured pri-/pre-miRNAs have the same potential to interact with target mRNAs, and it seems logic that evolution would utilize the intrinsic properties encoded in the structures as well as the sequences of miRNA precursors. Thus, it is logical to develop an inclusive framework that considers the potential of pri- and pre-miRNAs, as well as mature miRNAs, to function in target recognition and repression. Such a framework will allow experimental design, interpretations, and discussions that are balanced and inclusive.

### 3.1. Basic assumptions for the integrated framework

Here I outline a list of basic assumptions that I believe should serve as basis for the integrated and inclusive framework of target recognition and repression by the major RNA species produced from miRNA genes. Clearly outlined assumptions provide initial bases for further modifying and testing models derived from the framework.

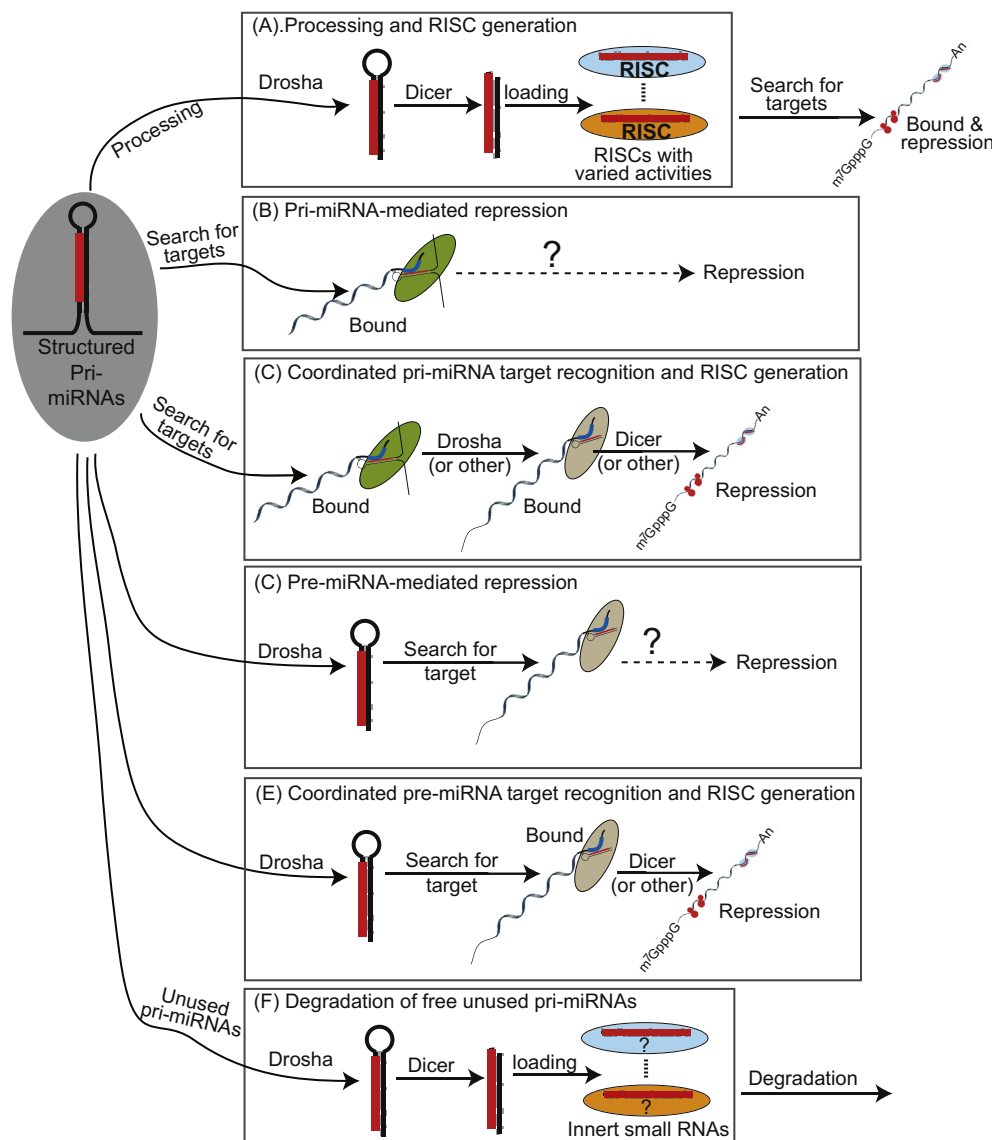
- The major products of an miRNA gene – pri-, pre-, and/or mature miRNAs – all have potential in target recognition and repression. Even if there was a gene for which the mature miRNA was shown to be the sole target-recognizing species, this finding could not be generalized to all miRNA genes.
- miRNA genes encode long primary transcripts that fold into structured RNAs containing stem–loop motifs. Structured pri-miRNAs and pre-miRNAs contain regulatory information not only in their primary sequences but also in their secondary and tertiary structural and sequence elements. These elements, absent from the mature miRNAs, are present in pri- and pre-miRNAs and can be used in target recognition and repression, as they are in bacterial antisense RNAs.
- Structured pri- and/or pre-miRNAs have the potential to interact with either linear or structured target RNAs. As illustrated by bacteria antisense RNAs, the interaction between structured pri-/pre-miRNAs and their target RNAs can be controlled by structural and sequence motifs, such as loop

nucleotides and stem–loops. The regulatory information encoded in structured regulatory RNAs can be directly translated into activity through interaction between pri-/pre-miRNAs and their cognate targets.

- There are branches at various stages of miRNA biogenesis *in vivo*. Pri- and/or pre-miRNAs may be processed through alternative pathways that deviate from the presumed linear miRNA biogenesis pathway, through which pri-miRNAs are processed into pre-miRNAs and then into mature miRNAs.
- Drosha and Dicer process many cellular RNAs in addition to precursor miRNAs. One cannot attribute the biological effects of *drosha* or *dicer* deletion to loss of a miRNA gene alone.

### 3.2. Modes of action of miRNA gene products

I propose an integrated framework for gene regulation by various RNAs produced from miRNA genes (Fig. 4). I consider that primary and precursor miRNAs, as well as mature miRNAs, all have



**Fig. 4.** Schematic diagrams depicting models of gene regulation by pri-/pre-/mature miRNAs and illustrating the distinctions in how information encoded in miRNA genes is translated into functions in target recognition and repression. (A) Canonical mature miRNA model of target recognition. (B) Pri-miRNA-guided target recognition that is independent of mature miRNA biogenesis. (C) Pri-miRNA-guided target recognition that acts in coordination in mature miRNA biogenesis. (D) Pre-miRNA-guided target recognition that is independent of mature miRNA biogenesis. (E) Pre-miRNA-guided target recognition that acts in coordination in mature miRNA biogenesis. (F) Degradation of unused pri-miRNAs through small RNA processing pathways. These pathways are not mutually exclusive and likely have intricate interrelationships.

the potential to interact with target miRNAs. This framework consists of six possible models of gene regulation (A–F) and considers the possible relationships among various miRNA gene products and their fates. Model A is the canonical miRNA model that depicts the mature miRNAs as the sole functional species in target recognition and repression. Model B is pri-miRNA-mediated target recognition and repression. In this model, pri-miRNAs directly bind to target mRNAs and mediate gene repression and do so independently of the canonical mature miRNA pathways. Model C describes coordinated target recognition and control of mature miRNA processing by pri-miRNAs. Under such a scenario, pri-miRNAs directly form complexes with target mRNAs, and pri-miRNAs are then processed into mature miRNAs by the canonical miRNA biogenesis machinery or other unknown ribonucleases. Many mature miRNA genes are produced at low copies in cells but are functional [59]. It is unclear how these low abundance mature miRNAs find their targets amongst abundant putative ones. In this model, the pri-miRNA and target interaction may help to guide low abundance mature miRNAs to selected targets. Model D is pre-miRNA-mediated target recognition and repression. In this mode, processed pre-miRNAs directly form complexes with target mRNAs, which are then processed by unknown ribonucleases independently of the canonical mature miRNA pathways. Model E is coordinated target recognition and control of mature miRNA processing by pre-miRNAs. In this mode, processed pre-miRNAs directly form complexes with target mRNAs, and pre-miRNAs are then processed by the canonical miRNA biogenesis machinery into mature miRNAs. Of note, pre-miRNAs, which form relatively simple stem–loop structures, may have weaker structural constraints than the unprocessed pri-miRNAs and more readily form duplexes with target RNAs. However, as illustrated by bacteria antisense regulatory RNAs, formation of fully duplexed complexes between antisense RNAs and their targets are not necessarily required for their regulatory functions. Thus, both pri- and pre-miRNA may regulate gene expression. Finally, model F depicts the turnover of non-functional pri-miRNAs. Under such a scenario, pri-miRNAs may be cleaved into non-functional mature miRNA products by the canonical miRNA biogenesis machinery or by other ribonucleases.

### 3.3. The relationships among the models

One of the most important distinctions among these models is how regulatory information encoded in the structured pri-miRNAs is translated into activity in target recognition and repression (Fig. 4). In model A, protein factors are evolved to interpret the information encoded in the structural and sequence elements of individual pri-miRNAs and translate it into activity through effects on biogenesis of mature miRNAs. In contrast, in models B, C, D, and E, the information encoded in the structured pri- and/or pre-miRNAs is directly translated into activity through interactions between structured RNAs. Model A is thus evolutionarily more costly since it requires active blockage of the use of intrinsic potential of structured RNAs to interact with other RNAs. Clearly, these distinctions determine the nucleotides and structural elements that should be utilized in experimental target identification and target prediction.

These models are not mutually exclusive and likely to have intricate relationships. For example, model F may have regulatory function by influencing the available functional species described in other models. Similarly, blocking pri-/pre-miRNAs from entering regulatory pathways may increase the fraction of pri-miRNAs that are non-productively processed in model F. Conversely, an increase in the rate of non-productive processing of pri-miRNAs in model F may reduce the availability of functional pri-/pre-miRNAs for target recognition and for functional mature miRNA biogenesis described in other models (A–E).

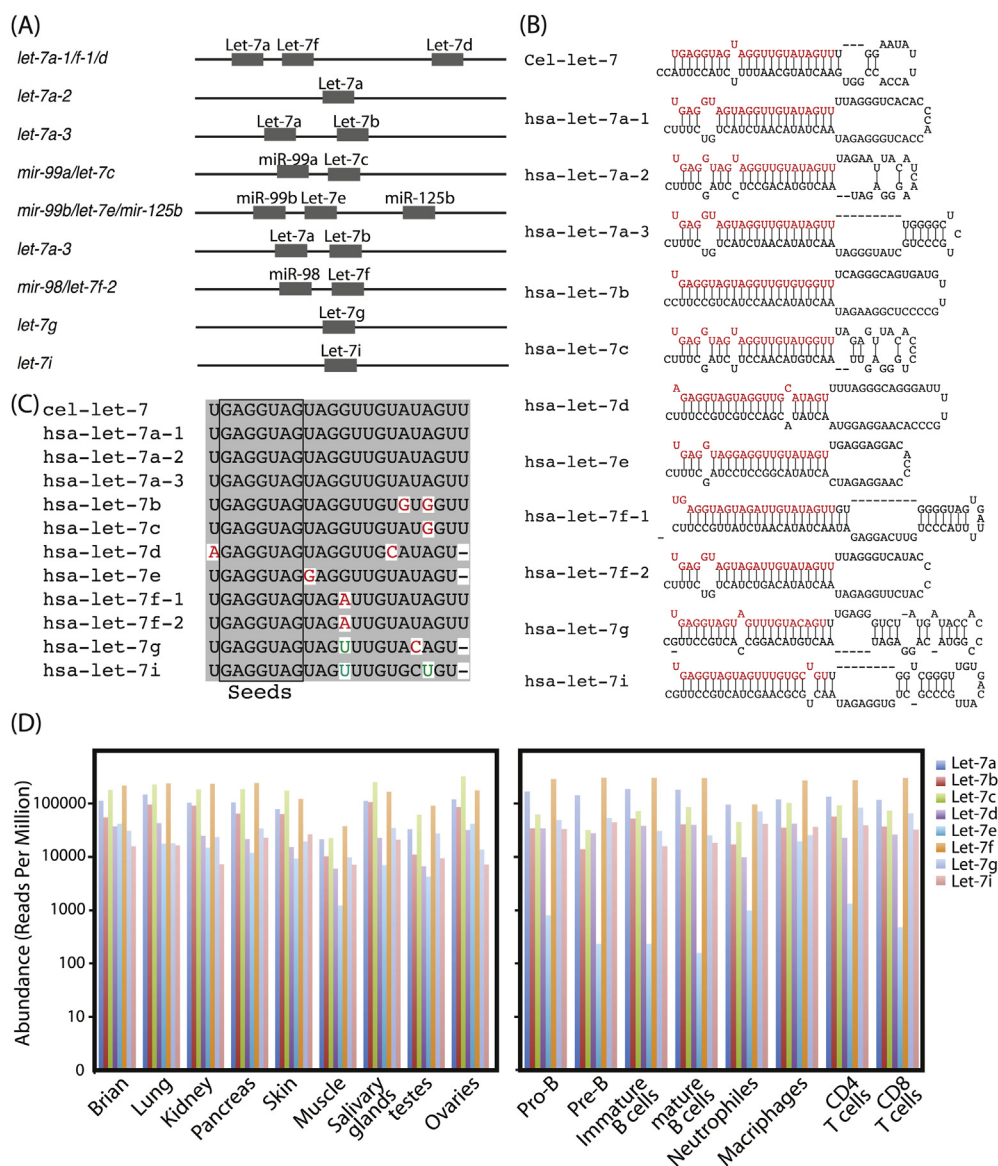
Importantly, one model does not have to fit all miRNA genes. It is possible that different pri-/pre-miRNAs may use distinct models of regulation and that the same pri-/pre-miRNAs may also operate through more than one mode of regulation for different targets. Specifically, one target mRNA may be regulated by the mature miRNA, whereas another may be regulated by the pri-miRNA/pre-miRNA encoded by the same gene. Finally, all models of regulation can be subjected to interference by proteins evolved to control critical regulatory steps. For example, inhibition of biogenesis of a mature miRNA would affect the fraction of pri-miRNAs that function through other models. In fact, protein factors that inhibit miRNA gene activities may do so by modulating mature miRNA biogenesis through interaction with pri-miRNAs or by blocking pri-/pre-miRNA:target interaction, thus there are alternative interpretations for some well-known observations if one considers a more inclusive framework [60–64].

## 4. Evidence in support of pri-/pre-miRNA function in target recognition and repression

The above analyses illustrate the complexity and challenge to solve the puzzle of which miRNA species play roles in target recognition and to what extent. However, with the proposed framework (Fig. 4), many important questions, testable hypotheses, and experimental solution are immediately apparent. In fact, many published studies have revealed curious results that beg for alternative explanations. Gatfield and colleagues showed that pri-miR-122 expression, which is regulated by circadian rhythm in mice, correlates with the changes in target expression, whereas mature miR-122 expression, which is not regulated by circadian rhythm and has a half-life longer than 24 h, does not correlate with the changes in target expression [65]. It has also been demonstrated that miRNAs encoded by members of the same family of genes have different biological activities [1,66], and that mature miRNAs made from artificial precursors are not functionally equivalent to the products of wild-type miRNA genes [67]. Mechanisms of miRNA-mediated translational repression are determined by the nuclear history of mRNA targets [68]. TRIM-NHL protein, NHL2, modulates the activity of let-7 miRNA genes without influencing mature miRNA levels [69]. Surprisingly, over 90% of mature miRNAs in mammalian cells are not in the RISCs; there is not enough Ago proteins to bind all available miRNAs [70]. Last but not the least, not only mature miRNAs but also pre-miRNAs and pri-miRNAs are found in RISC [71,72]. Collectively, all these findings require interpretations outside the canonical mature miRNA framework.

### 4.1. Evidence from the evolution of miRNA family genes

Dissecting the structural and sequence elements critical for the function of miRNA gene products can shed insight into the functions of each miRNA species without presumptions about the mechanisms of action by miRNA gene products. Interestingly, Nature has done the experiments for us through evolution. miRNA genes are often found as large families that encode highly homologous mature miRNAs but divergent pri-/pre-miRNA sequences. For example, there are eleven *let-7* miRNA genes that produce almost identical mature *let-7* miRNAs in the human genome (Fig. 5A–C). All *let-7* mature miRNAs have identical seed nucleotides (2–7 nucleotides of the 5' ends of mature miRNAs) and have one to three nucleotide variations outside of the seed regions. *let-7a-1*, *let-7a-2*, and *let-7a-3* produce identical mature *let-7a*, and *let-7f-1* and *let-7f-2* produce identical mature *let-7f*. However, the structures and sequences of the pri-/pre-miRNA stem–loop regions and precursor-flanking arms are significantly different. The evolution of miRNA genes through variations in their stem–loop regions is reminiscent



**Fig. 5.** The let-7 family genes and their overlapping mature miRNA expression patterns in mouse tissue and sorted hematopoietic cell types. (A) Schematics of the human let-7 gene family. (B) Secondary structure of human let-7 precursor miRNAs. (C) Alignment of human let-7 mature miRNAs. (D) The abundance of mature let-7 miRNA in mouse tissues and sorted hematopoietic cell types determined by analyzing RNA-seq results. Reads per million of specific mature let-7 miRNA species are shown.

to the evolution of bacteria antisense regulatory RNAs that control the plasmid copy numbers in bacteria (Figs. 3 and 5B). As I have illustrated, a few nucleotide differences in loop nucleotides or structural constraints can significantly alter activity and specificity of particular bacterial antisense RNAs even when perfect (or near perfect) matches to their target RNAs are retained (see Section 3). This remarkable parallel between the evolution of pri-miRNAs and bacterial antisense RNAs suggests that miRNA genes of the same family may have distinct biological activities under *in vivo* physiological conditions. Since each miRNA can regulate a large number of targets, the likely scenario is that miRNA genes producing identical mature miRNAs exert their distinct activities through regulating partially overlapping sets of target mRNAs that are determined by their differences in their pri-/pre-miRNAs.

One might argue that the duplication and diversification of miRNA genes have generated complex mature miRNA expression patterns through transcriptional and post-transcriptional regulation. However, members of the let-7 family miRNAs are often co-

expressed in the same tissues and cell types [73], suggesting that diversification of expression patterns is not the sole purpose of such duplication (Fig. 5D). A number of studies have shown that protein factors bind to the loop nucleotides of pri-/pre-miRNAs, and some these factors control the biogenesis of mature miRNAs in some cell lines [74]. However, in the inclusive framework, these observations can also be explained by the pri-/pre-miRNAs function in target regulation (Fig. 4, models B, C, D, E). These proteins may block the interactions between pri-/pre-miRNAs and their target RNAs by masking the recognition motifs, such as the loop sequences. Unused pri-/pre-miRNAs may be then subjected to non-productive turnover (Fig. 4, models F and others). It is also conceivable that some proteins may have evolved to potentiate interactions between pri-/pre-miRNAs and mRNAs by stabilizing these complexes. If one considers that only mature miRNA are functional, it means that evolution must purposefully reject the use of structured RNAs as regulatory molecules and develop a system to actively prevent the interaction between structured miRNA precursors and their

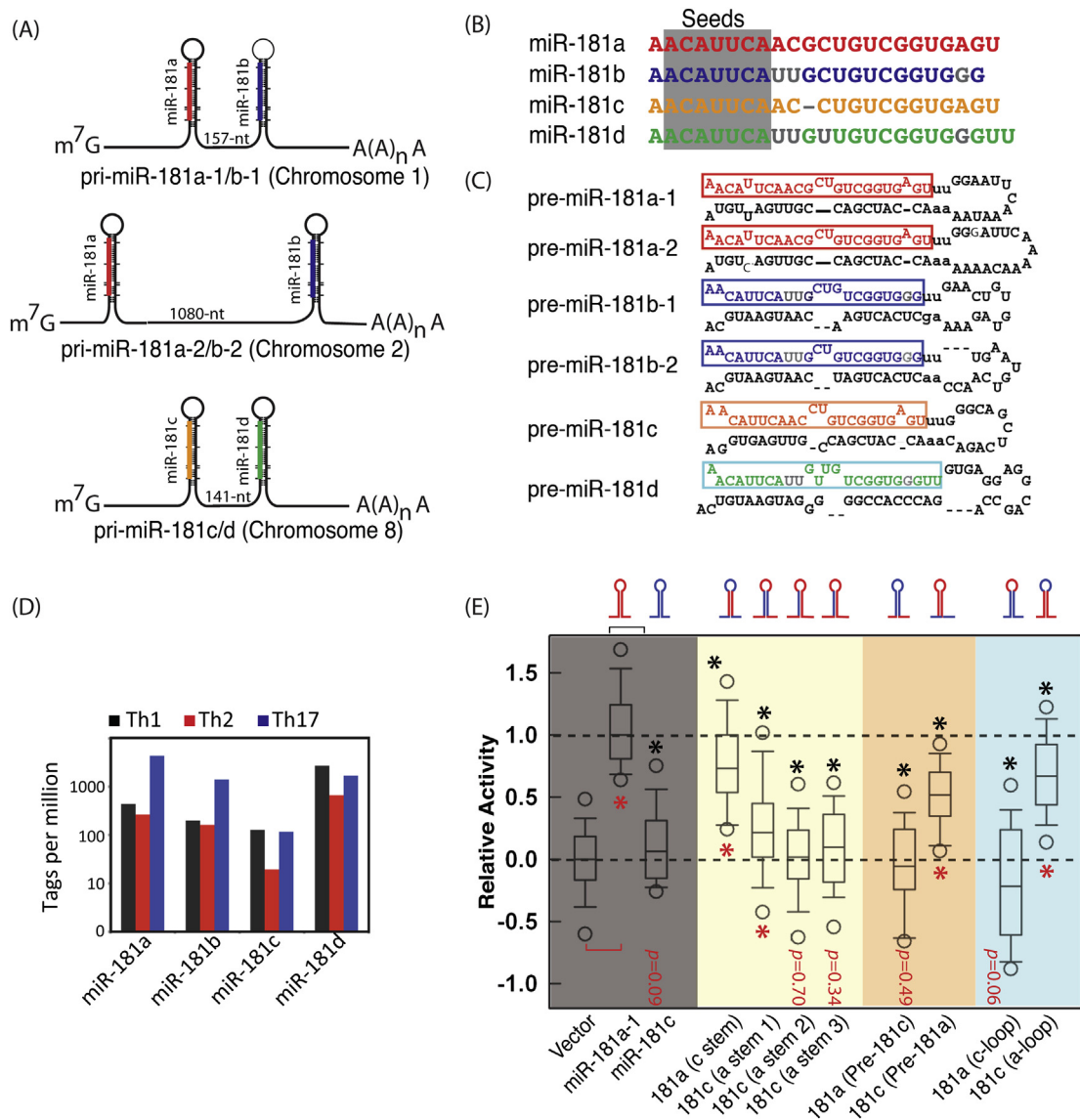


cognate targets. Although this is possible, the benefits of such an evolutionary choice are unclear. Collectively, given that it is the intrinsic molecular property of RNA to form structures that mediate interactions with RNAs and other biological molecules, it is reasonable to propose that structural properties of pri-/pre-miRNAs are used for target recognition and repression.

#### 4.2. Loop nucleotides control the distinct activities of miRNA genes

Do miRNA genes that produce identical or nearly identical miRNAs but different loop nucleotides have the same biological activities? If pri-miRNAs/pre-miRNAs work through models B–E (Fig. 4), then pri-/pre-miRNAs that encode identical mature miRNAs but that have distinct structures and sequence motifs (i.e., stem–loops and flanking sequences) may have different biological activities. The members of the miR-181 family of genes produce four nearly

identical mature miRNAs – miR-181a, miR-181b, miR-181c, and miR-181d – from three polycistronic miRNA genes – *mir-181a-1/b-1*, *mir-181a-2/b-2*, and *mir-181c/d* (Fig. 6A–C). The miR-181 miRNAs are co-expressed by many T cell types (Fig. 6D), albeit at varied levels [73]. We asked whether *mir-181a-1* and *mir-181c*, which produce mature miRNAs with only a single nucleotide difference in the center of mature miRNAs, have distinct functions in early T cell development [1]. We showed that ectopic expression of *mir-181a-1* potentiates the generation of CD4 and CD8 double-positive (DP) cells in an *in vitro* culture assay (Fig. 6E). Interestingly, *mir-181a-1*, but not *mir-181c*, can promote DP T cell development when ectopically expressed in thymic progenitor cells. Importantly, the distinct activities of *mir-181a-1* and *mir-181c* are largely determined by their unique pre-miRNA loop nucleotides – not by the single nucleotide differences in their mature miRNA sequences. Moreover, the effects of *mir-181a-1* on DP cell development are quantitatively influenced



**Fig. 6.** The miRNA family genes and their differential function controlled by their unique loop sequences. (A) Schematics of mouse miR-181 gene family. (B) Secondary structure of miR-181 precursor miRNAs. (C) Alignment of miR-181a mature miRNAs. (D) The abundance of mature miR-181 miRNA in T helper cell types, Th1, Th2, and Th17, determined by analyzing RNA-seq results. Reads per million of specific mature let-7 miRNA species are shown. (E) The pri-/pre-miRNA loop nucleotides control the distinct activities of *mir-181a-1* and *mir-181c*. Schematics depicting various *mir-181a-1* and *mir-181c* domain-swapping mutants and their effects on DP cell development [1]. Normalized data from three to seven independent T cell assays (each with 12 independent replicates for a total of 36–84 replicates) were pooled and graphed in the distribution box plots. Significance was determined by Mann–Whitney rank sum tests (compared to the negative control vector, red, \**p* < 0.0001; compared to the *mir-181a-1* positive control, black, \**p* < 0.0001).

by nucleotide changes in the pri-/pre-miRNA loop nucleotide region; both the strength and the functional specificity of miR-181 genes depend on the pri-/pre-miRNA loop sequences [1]. Mutations in the pre-miRNA loop regions affect pre-miRNA and mature miRNA processing, but we found no consistent correlation between the levels of mature miRNAs and the activities of the *mir-181a-1/c* genes with loop nucleotide mutations. These results demonstrate that miRNA genes producing identical or nearly identical mature miRNAs have distinct biological activities resulting from distinct pri-/pre-miRNA loop sequences, supporting the hypothesis that miRNA genes likely evolved to achieve different activities via alterations in their pri-/pre-miRNA loop sequences without altering mature miRNA sequences.

The relationships between mature miRNA levels and activities are not the same for each of the models shown in Fig. 4. For the canonical model A, miRNA genes encoding identical mature miRNAs should have the same activity when these mature miRNAs are co-expressed in the same cells. Thus, the distinct activities of miRNA family genes can be only explained by differential mature miRNA biogenesis. In contrast, in models B, C, D, and E, miRNA genes encoding identical mature miRNAs can have different activities even when the mature miRNAs are co-expressed in the same cells. Moreover, mature miRNA levels and activities should correlate for models C and E but should not necessarily correlate for models B and D. Since a significant fraction of mature miRNAs is not associated with RISC [70], pri-/pre-miRNAs that are not being used for regulation may be degraded by RNase III enzymes into small ~22-nt RNAs as suggested by model F. If a large fraction of pri-/pre-miRNAs is processed into non-functional miRNAs, these non-functional mature miRNAs may obscure the true correlation between mature miRNA levels and activities. Therefore, cautions must be taken to use the relationship between mature miRNA levels and activities to extrapolate the potential functional species of miRNA genes.

The interpretation of our *mir-181* gene findings clearly depends on the model one adopts. If one assumes that only model A is correct, one has to assume that loop mutations cause differential loading of mature miRNAs onto RISC or that these mutations alter rates of processing and therefore activity; these assumptions are difficult if not impossible to test experimentally. If one is open to other models, these results suggest that either pri-miRNAs or pre-miRNAs have direct roles in target recognition and repression and loop nucleotides control or modulate these interactions [2,3]. Since swapping pre-miRNAs or loop regions between *mir-181a-1* and *mir-181c* does not fully exchange activity, it is likely that pri-miRNAs contain all elements required for the target regulation and therefore are functional. Our results demonstrate that loop nucleotides play critical roles in controlling activities of pri-/pre-miRNAs, as stem-loops would do in controlling the activity of bacterial antisense RNAs, lending strong experimental support to the remarkable parallels between pri-/pre-miRNAs and bacteria antisense RNAs [1].

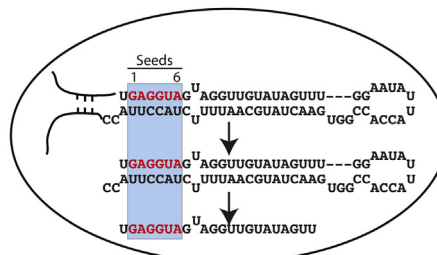
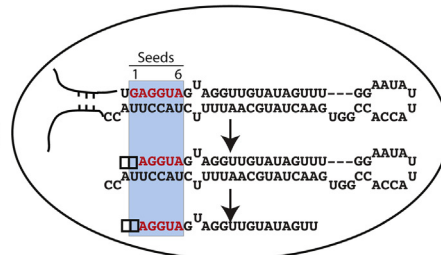
Is this phenomenon specific to *mir-181a-1* and *mir-181c* during early T cell development? Recently, through loss of function analyses, we demonstrated that deletion of *mir-181a-1/b-1*, but not *mir-181-c/d* or *mir-181a-2/b-2*, selectively inhibits tumor transformation induced by Notch oncogenes [66]. Moreover, deletion of *mir-181a-2/b-2*, but not *mir-181a-1/b-1* and *mir-181-c/d*, affects the rate of embryonic stem cells (ESCs) self-renewal despite the fact that all three alleles are expressed in ESCs (Arnold et al., unpublished observations). It is important to note that *mir-181a-2/b-2* and *mir-181a-1/b-1* produce identical mature miR-181a and miR-181b. Finally, loop nucleotides also play critical roles in controlling target repression by let-7 miRNA genes [2,3]. Thus, target recognition by the structured pri-/pre-miRNA species may be utilized as a mechanism in target recognition by many miRNA genes.

These findings demonstrate that miRNA genes encoding identical or nearly identical mature miRNAs can have distinct biological activities despite their overlapping mature miRNA expression patterns and that their differential effects are independent of their mature miRNA levels. It is possible to use similar biologically relevant systems to dissect differential effects of the miRNA gene family members on mRNA and protein expression at global levels. One may be able to identify a distinct set or a partially overlapping set of targets that are controlled by miRNA genes that produce identical and or nearly identical mature miRNAs under physiologically relevant settings. Identification of distinct target sets that are determined by pri-miRNAs would further support the role of pri-/pre-miRNAs in target recognition. With the identification of such targets, it would be possible to further investigate whether pri-miRNA loops play direct roles in base-pairing recognition of cognate targets.

#### 4.3. Independent function of pri-miRNAs in target recognition and repression

Is there a way to dissect the functions of pri-miRNAs (or pre-miRNAs) in target recognition and repression independent of mature and pre-miRNA? In my opinion, it is critical to develop a system to examine the function of pri-miRNAs under normal cellular conditions and without the complications of *dicer* and *droscha* null cells. To this end, we have fortuitously discovered a condition that generates pri-miRNAs but not functional pre-miRNAs and mature miRNAs [2]. When we expressed *C. elegans let-7* gene (*cel-let-7*) in human cells, we found that *cel-let-7*, which is capable of repressing the expression of luciferase reporters with the *lin-41* target 3' UTR, produces pre- and mature let-7 miRNAs that lack two nucleotides on the 5' ends (Fig. 7A and B). These pre- and mature let-7 miRNAs missing the first seed nucleotides are not functional in target repression. Moreover, despite that primary *cel-let-7* is capable of target repression in the absence of functional pre- and mature let-7, its activity is dependent on the first seed nucleotide that presents in primary *cel-let-7* but is missing in pre- and mature let-7 (Fig. 7C). Moreover, we showed that *cel-let-7* pri-miRNAs form complexes with target mRNAs containing the *lin-41* 3' UTR *in vitro* in the absence of RISC or any other proteins and *in vivo*. Importantly, complex formation between pri-*cel-let-7* and *lin-41* 3' UTR is controlled by the seed and loop nucleotides and the ability of these mutant forms of these two RNAs to form complexes correlates with the extent target repression, indicating a direct role of pri-let-7a loop in the formation of pri-miRNA:target complexes. However, whether pri-let-7a forms complexes with its target mRNA through kissing-loop intermediates that resemble those of bacteria antisense RNAs remains to be established. Finally, we showed that human *let-7a-3* pri-miRNA also forms complexes with target mRNAs. Together, these results provided strong evidence that primary miRNAs can have direct roles in target recognition and repression. Our findings suggest that the ancient mechanisms bacteria antisense RNA inherited from the RNA World are alive and well for the miRNA genes, and that the regulatory information encoded in the structural and sequence elements of pri-miRNAs can be directly translated into activities in target recognition and repression through interactions with their target mRNAs.

Could it be possible that, if mature miRNAs were properly generated, primary miRNAs would not be needed and the function of miRNA genes will then be entirely executed by mature miRNAs? To this end, we showed that modulating pri-let-7 loop nucleotides restored the two missing nucleotides at the 5' end of mature and pre-let-7 and significantly increased the activity of *cel-let-7* gene [3]. We found that human *let-7a-3* pri-miRNAs formed complexes with target mRNAs even in the presence of full-length and

(A). *Cel-let-7* expression in worms(B). *Cel-let-7* expression in BOSC23 cells(C). *Cel-let-7* expression in worms

<i>cel-let-7</i> variations	Full-length Mature	Relative In vivo pri-miRNA:target Complex	Repression activity (%)
<i>cel-let-7_wt</i>	No	1	38.7±4.87
<i>cel-let-7_2A</i>	No	—	0
<i>cel-let-7_LPA</i>	No	0.42±0.06	13.6±6.7
<i>cel-let-7_LPB</i>	No	1.2±0.1	48.9±4.1
<i>cel-let-7_LPC</i>	No	0.29±0.06	0
<i>cel-let-7_LPD</i>	No	0.94±0.22	26.4±6.6
<i>cel-let-7_LP4</i>	Yes	0.84±0.22	78±4.0

**Fig. 7.** Target recognition and repression by pri-miRNAs in the absence of functional mature miRNAs. (A) The 5' ends of pre-let-7 and mature let-7 made in *C. elegans*. (B) The pre-let-7 and mature let-7 made in BOSC 23 cells, a derivative of primary embryonic human kidney 293 cells, lack two nucleotides at their 5' ends including the first seed nucleotide of let-7 miRNA. (C) Function of *cel-let-7* mutants in target recognition and repression in the presence of truncated and non-functional mature let-7 miRNAs. The tables summarize the ability of the *cel-let-7* mutants to produce functional mature miRNAs, form pri-let-7 and target complexes, and suppress luciferase target reporters [2,3].

theoretically functional mature let-7 [2]. Thus, even with the presence of full-length mature let-7, *cel-let-7* pri-miRNA still forms complexes with target mRNAs. Interestingly, pri- and mature miRNA contributions to target repression could be distinguished by the differential sensitivity to the pairing to the first seed nucleotide [3]. We noted that some of the null-loop mutants described in the Trujillo study (e.g., LP-C, Fig. 7C) rescue the *n2853* temperature sensitive let-7 mutant (Trujillo & Chen, unpublished observation, and Zhang and Fire, personal communication). It is necessary to further dissect the function of endogenous pri-miRNAs in target recognition and repression in the presence of mature miRNAs; this may be facilitated by identification of processing enzymes and other regulatory proteins.

## 5. Questions and testable hypotheses

The studies described above strongly support the hypothesis that pri-miRNAs (and/or pre-miRNAs) play direct roles in target recognition and repression. However, it remains challenging to determine to what extent mature and pri-/pre-miRNA species contribute to gene repression under physiological conditions. It is also difficult to determine whether pri-, pre-, and mature miRNAs function independently or coordinately if multiple species are important. The integrated models presented here (Fig. 4) provide a framework for experimental design and interpretation that may be key to resolve the puzzle of functional miRNA species (Fig. 1). Here I will briefly discuss a few questions suggested by the models (Fig. 4).

### 5.1. Branches in miRNA biogenesis pathway

The existence of branches in the miRNA biogenesis pathway [74] are consistent with the framework proposed (Fig. 4). For example, a

number of proteins, including KSSRP, Lin-28, and HRNAP A1, have been shown to bind to selected sets of pri-miRNAs or pre-miRNAs and divert them away from the canonical miRNA biogenesis [60–64]. These pri-miRNAs and pre-miRNAs are presumably directed into degradation pathways, though their exact fates remain unclear. It is important to note that the effects of these proteins on the activities of miRNA genes can be explained by their ability to reduce the availability of structured pri-/pre-miRNAs for target recognition (models B–E). Branches may also exist in other steps of miRNA biogenesis. In an elegant recent study, Janas and colleagues showed that only ~10% mature miRNAs are actually associated with Ago proteins in HeLa cells, demonstrating for the first time that most cellular mature miRNAs are not in a RISC. This observation revealed an unexpected branch in mature miRNA biogenesis *in vivo* [70], suggesting a cautionary note for direct extrapolation of *in vitro* data to *in vivo* mechanisms. Also important to note, mature miRNAs may be loaded into various forms of RISC that contain different Argonaute proteins (i.e., Ago 1–4 in human cells), but the fractions of mature miRNAs that are loaded into particular Ago-RISCs *in vivo* and their functional significances are unknown. More importantly, if the majority of mature miRNAs are not associated with RISCs, what is the function of these RISC-free mature miRNAs *in vivo*?

The models presented here (Fig. 4) suggest that mature miRNAs may be generated from various pathways and may have distinct functionalities and/or destinations. miRNAs are generated and load into RISCs in model A, and non-functional miRNAs are produced in model F. In contrast, the mature miRNAs generated in models B and D are already bound to targets, whereas mature miRNAs may or may not be generated in models C and E. The double-stranded pri-miRNA:target complexes generated in models B–E can be degraded either by RNase III enzymes that produce ~22-nt small RNAs or by

other ribonucleases that do not. It is also of interest to determine whether miRNAs generated through different pathways are actually loaded into different RISC complexes or have different modifications.

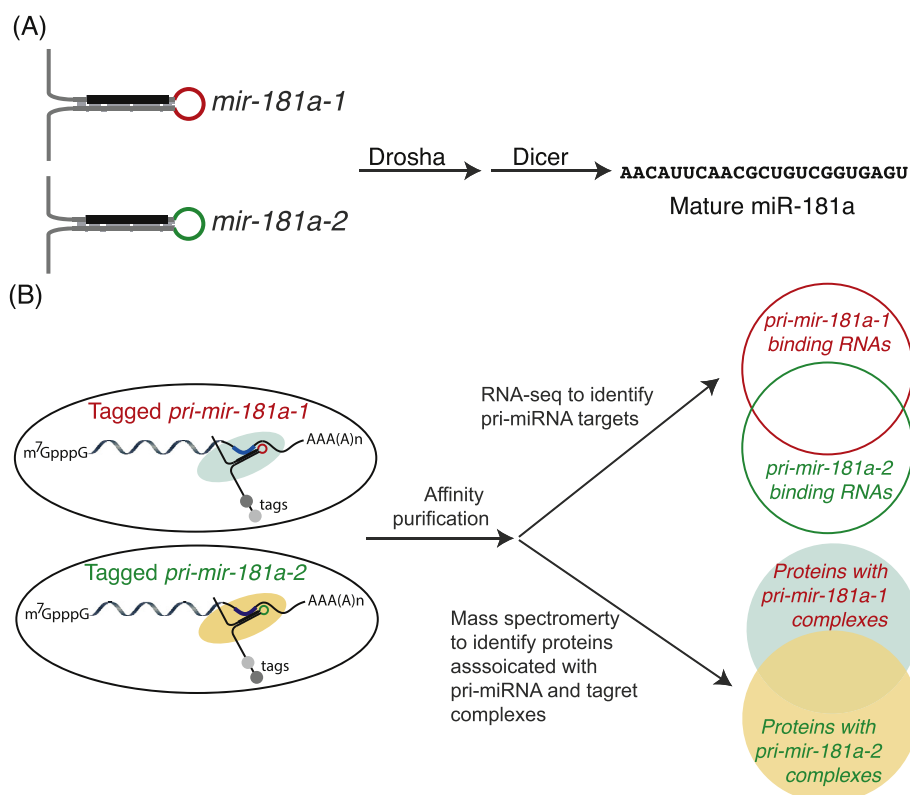
The observed branches in the mature miRNA processing pathway raise questions regarding the canonical miRNA biogenesis model *in vivo*. Given the intricate connections among various species of miRNA gene products, inhibition of miRNA processing through model A or model E may increase the fraction of pri-miRNA and pre-miRNAs available for gene regulation through models B, C, D, and E. Moreover, recent studies showed that mature miRNAs might be stabilized by the presence of target mRNA [75]. Thus, a quantitative model is needed to dissect how various regulatory events affect the processing and activities of pri-miRNA, pre-miRNA, and mature miRNA. Quantitative models may be used to discern how changes of target levels affect the stability of pri-miRNA, pre-miRNA, and mature miRNA species *in vivo*. Distinct outcomes may be postulated depending on the model of regulation. For example, pri-miRNAs (model B and C) or pre-miRNAs (model D and E) in the target complexes may have different stabilities from those free pri-miRNAs (model A and F). Another intriguing possibility is that unprocessed or partially processed pri-miRNAs may be present and active in cytoplasm. A few primary miRNAs are found in the cytoplasm [76], suggesting that there is a cytoplasmic pathway for pri-miRNA processing or activity.

## 5.2. Target identification using pri-miRNAs as baits

Assuming that pri-miRNAs directly bind target mRNAs (Fig. 4, models B and C), it should be possible to identify target mRNAs by using pri-miRNAs as bait. Such an approach should be feasible since a target mRNA tagged with S1 aptamer, which binds to streptavidin

beads, was successfully used to pull-down the target and pri-miRNA complexes [2,3]. Therefore, one can tag a pri-miRNA of interests with S1 or other affinity tags, purify mRNAs associated with the tagged pri-miRNAs, and then determine the identities of mRNAs through RNA-seq analyses and proteins through mass spectrometry analyses (Fig. 8). This approach can be used to examine whether the pri-miRNAs producing identical mature miRNAs, such as *mir-181a-1* and *mir-181a-2*, recognize shared or non-overlapping target sets (Fig. 8). Moreover, it would be of interest to compare target mRNAs enriched by pri-miRNAs to those enriched by the corresponding mature miRNAs via HIT-CLIP [77]. Such comparisons will help determine whether pri-miRNAs and mature miRNAs recognize shared or non-overlapping target sets, providing insight into the relationships between pri-miRNA- and mature miRNA-mediated target regulation. Finally, target sequences generated from such analyses can be used to decipher sequence and structural elements of pri-miRNAs important for pri-miRNA and target mRNA complex formation. Such information may reveal principles underlying target recognition that could be incorporated into computational programs for target prediction to further improve these algorithms.

By analogy to bacterial antisense RNAs (Fig. 3), protein factors are likely to play critical roles in facilitating the complex formation between pri-miRNAs and their mRNAs. Moreover, different pri-miRNAs may bind distinct protein complexes. The pri-miRNA pull-down strategy described above may be employed to identify protein factors associated with pri-miRNA:target complexes. Proteomic analyses can be carried out identify the proteins associated with pri-miRNA:target complexes, and their functions in pri-miRNA-mediated repression can be dissected through gain-of- and loss-of-function analyses. Intriguingly, we showed that pri-let-7 and target complexes were found in the nucleus, suggesting



**Fig. 8.** Target identification through pri-miRNA pull-down analyses. (A) Schematics depicting pri-*mir-181a-1* and pri-*mir-181a-2*, which produce identical mature miR-181a, but that have distinct hairpin loop sequences. (B) Schematics depicting a strategy to identify pri-*mir-181a-1* and pri-*mir-181a-2* binding mRNAs and proteins. Pri-miRNAs tagged with RNA aptamer (i.e., S1-Tag) can be used to enrich pri-miRNA and mRNA complexes and their associated proteins. RNAs and proteins can be identified by RNA-seq and mass spectrometry analyses, respectively. It is possible that pri-*mir-181a-1* and pri-*mir-181a-2* bind to partially overlapping sets of mRNAs and proteins.



that pri-miRNA-mediated gene regulation may be initiated in the nucleus [2]. However, pri-let-7 and target complexes are also observed in the cytoplasm [2], and the protein machinery involved in the degradation of such complexes and the location of the degradation remain unclear. Systematic analyses of protein factors critical for pri-miRNA-mediated target regulation will shed light on the mechanisms of repression by pri-miRNAs and will reveal the differences between the mechanisms of repression by pri-miRNAs and mature miRNAs.

### 5.3. Perspectives and concluding remarks

In summary, the proposed framework (Fig. 4) integrates the functions of pri-/pre-miRNAs and mature miRNAs in target recognition and repression and will open areas of experimental enquiry and permit interpretations precluded by the canonical mature miRNA model. There is no doubt that the discovery of miRNA genes represents a major milestone in modern molecular biology, and genetic analyses have revealed fascinating insights into the roles of miRNA genes. However, there are still many unanswered questions and controversies with regard to target recognition and mechanisms of repression by the miRNA gene products. Not knowing the exact contributions of mature and pri-/pre-miRNAs in target recognition may be at the root of many controversies.

The problems with the mature miRNA assumption are easily discerned by answering following questions: (1) Is there a published study that established the mature miRNAs as the sole functional species of miRNA gene products? (2) Can we generalize the mature miRNA assumption from one example to all miRNA genes? (3) Is it an intrinsic property of RNA molecules to form secondary and tertiary structures? (4) Do examples from the RNA World clearly illustrate that the regulatory information encoded in the structured RNA molecules can be directly translated into activity through interactions between structured RNA molecules? (5) Is there any convincing evidence supporting that miRNA genes have evolved to avoid utilizing the intrinsic properties of structured pri-/pre-miRNA to interact with other RNAs? Although what we have learned so far about mature miRNAs likely represents an important facet of miRNA biology, it is important to be open-minded about potentially diverse mechanisms through which pri- and/or pre-miRNAs contribute to target recognition and repression. This article represents a preliminary effort to conceptualize the interesting processes that are controlled by pri- and/or pre-miRNAs and the relationships of these processes with mature miRNAs. I hope that the analyses will facilitate more inclusive analyses of the mechanisms of action by the various RNAs produced from miRNA genes.

### Acknowledgments

I would like to thank Drs. Shenglin Huang, Christina Loh, Vlad Elgart, Rita Fragoso, and Andrew Fire for their critical comments on the manuscript. This work was supported by W.M. Keck Young Scholar Award and NIH grants (1R01AI073724 and 1DP1 OD00643501) to C.-Z. C.

### References

- [1] G. Liu, H. Min, S. Yue, C.Z. Chen, Pre-miRNA loop nucleotides control the distinct activities of mir-181a-1 and mir-181c in early T cell development, *PLoS ONE* 3 (2008) e3592. <http://dx.doi.org/10.1371/journal.pone.0003592>.
- [2] R.D. Trujillo, S.B. Yue, Y. Tang, W.E. O'Gorman, C.Z. Chen, The potential functions of primary microRNAs in target recognition and repression, *EMBO J.* 29 (2010) 3272–3285. <http://dx.doi.org/10.1038/emboj.2010.208>.
- [3] S.B. Yue, R.D. Trujillo, Y. Tang, W.E. O'Gorman, C.Z. Chen, Loop nucleotides control primary and mature miRNA function in target recognition and repression, *RNA Biol.* (2011) 1115–1123. <http://dx.doi.org/10.4161/rna.8.6.17626>.
- [4] N.C. Lau, L.P. Lim, E.G. Weinstein, D.P. Bartel, An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, *Science* 294 (2001) 858–862.
- [5] R.C. Lee, V. Ambros, An extensive class of small RNAs in *Caenorhabditis elegans*, *Science* 294 (2001) 862–864.
- [6] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, Identification of novel genes coding for small expressed RNAs, *Science* 294 (2001) 853–858.
- [7] D.P. Bartel, MicroRNAs: target recognition and regulatory functions, *Cell* 136 (2009) 215–233.
- [8] V.N. Kim, J. Han, M.C. Siomi, Biogenesis of small RNAs in animals, *Nat. Rev. Mol. Cell Biol.* 10 (2009) 126–139.
- [9] V. Ambros, The evolution of our thinking about microRNAs, *Nat. Med.* 14 (2008) 1036–1040.
- [10] S.M. Elbashir, W. Lendeckel, T. Tuschl, RNA interference is mediated by 21- and 22-nucleotide RNAs, *Genes Dev.* 15 (2001) 188–200.
- [11] W. Filipowicz, S.N. Bhattacharyya, N. Sonenberg, Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9 (2008) 102–114.
- [12] D. Didiano, O. Hobert, Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions, *Nat. Struct. Mol. Biol.* 13 (2006) 849–851.
- [13] D. Didiano, O. Hobert, Molecular architecture of a miRNA-regulated 3' UTR, *RNA* 14 (2008) 1297–1317.
- [14] M. Thomas, J. Lieberman, A. Lal, Desperately seeking microRNA targets, *Nat. Struct. Mol. Biol.* 17 (2010) 1169–1174.
- [15] R.C. Lee, R.L. Feinbaum, V. Ambros, The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*, *Cell* 75 (1993) 843–854.
- [16] A. Fire, S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, C.C. Mello, Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature* 391 (1998) 806–811.
- [17] A.J. Hamilton, D.C. Baulcombe, A species of small antisense RNA in post-transcriptional gene silencing in plants, *Science* 286 (1999) 950–952.
- [18] P.D. Zamore, T. Tuschl, P.A. Sharp, D.P. Bartel, RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals, *Cell* 101 (2000) 25–33.
- [19] T. Tuschl, P.D. Zamore, R. Lehmann, D.P. Bartel, P.A. Sharp, Targeted mRNA degradation by double-stranded RNA *in vitro*, *Genes Dev.* 13 (1999) 3191–3197.
- [20] J.G. Doench, C.P. Peterson, P.A. Sharp, siRNAs can function as miRNAs, *Genes Dev.* 17 (2003) 438–442.
- [21] E. Bernstein, A.A. Caudy, S.M. Hammond, G.J. Hannon, Role for a bidentate ribonuclease in the initiation step of RNA interference, *Nature* 409 (2001) 363–366.
- [22] A. Grishok, A.E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D.L. Baillie, A. Fire, G. Ruvkun, C.C. Mello, Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing, *Cell* 106 (2001) 23–34.
- [23] S.M. Hammond, S. Boettcher, A.C. Caudy, R. Kobayashi, G.J. Hannon, Argonaute 2, a link between genetic and biochemical analyses of RNAi, *Science* 293 (2001) 1146–1150.
- [24] G. Hutvagner, P.D. Zamore, A microRNA in a multiple-turnover RNAi enzyme complex, *Science* 297 (2002) 2056–2060.
- [25] A.J. Giraldez, R.M. Cinalli, M.E. Glasner, A.J. Enright, J.M. Thomson, S. Baskerville, S.M. Hammond, D.P. Bartel, A.F. Schier, MicroRNAs regulate brain morphogenesis in zebrafish, *Science* 308 (2005) 833–838.
- [26] C. Melton, R.L. Judson, R. Blueloch, Opposing microRNA families regulate self-renewal in mouse embryonic stem cells, *Nature* 463 (2010) 621–626.
- [27] J. Han, J.S. Pedersen, S.C. Kwon, C.D. Belair, Y.K. Kim, K.H. Yeom, W.Y. Yang, D. Haussler, R. Blueloch, V.N. Kim, Posttranscriptional crossregulation between Drosha and DGCR8, *Cell* 136 (2009) 75–84.
- [28] S. Kadener, J. Rodriguez, K.C. Abruzzi, Y.L. Khodor, K. Sugino, M.T. Marr 2nd, S. Nelson, M. Rosbash, Genome-wide identification of targets of the drosha-pasha/DGCR8 complex, *RNA* 15 (2009) 537–545.
- [29] A. Nakagawa, Y. Shi, E. Kage-Nakadai, S. Mitani, D. Xue, Caspase-dependent conversion of dicer ribonuclease into a death-promoting deoxyribonuclease, *Science* 328 (2010) 327–334.
- [30] B.J. Reinhart, F.J. Slack, M. Basson, J.C. Bettinger, A.E. Pasquinelli, A.E. Rougvie, H.R. Horvitz, G. Ruvkun, The 21 nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*, *Nature* 403 (2000) 901–906.
- [31] Y. Eguchi, T. Itoh, J. Tomizawa, Antisense RNA, *Annu. Rev. Biochem.* 60 (1991) 631–652.
- [32] R.W. Simons, N. Kleckner, Biological regulation by antisense RNA in prokaryotes, *Annu. Rev. Genet.* 22 (1988) 567–600.
- [33] C. Brunel, R. Marquet, P. Romby, C. Ehresmann, RNA loop-loop interactions as dynamic functional motifs, *Biochimie* 84 (2002) 925–944.
- [34] S. Gottesman, G. Storz, Bacterial small RNA regulators: versatile roles and rapidly evolving variations, *Cold Spring Harbor Perspect. Biol.* 3 (2011). pii: a003798. <http://dx.doi.org/10.1101/cshperspect.a003798>.
- [35] H. Grosjean, D.G. Soll, D.M. Crothers, Studies of the complex between transfer RNAs with complementary anticodons. I. Origins of enhanced affinity between complementary triplets, *J. Mol. Biol.* 103 (1976) 499–519.
- [36] J. Eisinger, N. Gross, The anticodon–anticodon complex, *J. Mol. Biol.* 88 (1974) 165–174.

- [37] Y. Eguchi, J. Tomizawa, Complexes formed by complementary RNA stem-loops. Their formations, structures and interaction with ColE1 Rom protein, *J. Mol. Biol.* 220 (1991) 831–842.
- [38] K. Okamoto, M. Freundlich, Mechanism for the autogenous control of the crp operon: transcriptional inhibition by a divergent RNA transcript, *Proc. Natl. Acad. Sci. U. S. A.* 83 (1986) 5000–5004.
- [39] J.C. Paillart, R. Marquet, E. Skripkin, B. Ehresmann, C. Ehresmann, Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA, *J. Biol. Chem.* 269 (1994) 27486–27493.
- [40] T. Maniatis, R. Reed, The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing, *Nature* 325 (1987) 673–678.
- [41] K. Nordstrom, E.G. Wagner, Kinetic aspects of control of plasmid replication by antisense RNA, *Trends Biochem. Sci.* 19 (1994) 294–300.
- [42] J. Tomizawa, Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript, *Cell* 38 (1984) 861–870.
- [43] J. Tomizawa, Control of ColE1 plasmid replication: initial interaction of RNA I and the primer transcript is reversible, *Cell* 40 (1985) 527–535.
- [44] J. Tomizawa, T. Som, Control of ColE1 plasmid replication: enhancement of binding of RNA I to the primer transcript by the Rom protein, *Cell* 38 (1984) 871–878.
- [45] M. Brenner, J. Tomizawa, Quantitation of ColE1-encoded replication elements, *Proc. Natl. Acad. Sci. U. S. A.* 88 (1991) 405–409.
- [46] R.P. Novick, Plasmid incompatibility, *Microbiol. Rev.* 51 (1987) 381–395.
- [47] J. Tomizawa, T. Itoh, Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript, *Proc. Natl. Acad. Sci. U. S. A.* 78 (1981) 6096–6100.
- [48] E.G. Wagner, R.W. Simons, Antisense RNA control in bacteria, phages, and plasmids, *Annu. Rev. Microbiol.* 48 (1994) 713–742.
- [49] C. Persson, E.G. Wagner, K. Nordstrom, Control of replication of plasmid R1: formation of an initial transient complex is rate-limiting for antisense RNA–target RNA pairing, *EMBO J.* 9 (1990) 3777–3785.
- [50] C. Persson, E.G. Wagner, K. Nordstrom, Control of replication of plasmid R1: structures and sequences of the antisense RNA, CopA, required for its binding to the target RNA, CopT, *EMBO J.* 9 (1990) 3767–3775.
- [51] K.M. Wassarman, A. Zhang, G. Storz, Small RNAs in *Escherichia coli*, *Trends Microbiol.* 7 (1999) 37–45.
- [52] S. Altuvia, D. Weinstein-Fischer, A. Zhang, L. Postow, G. Storz, A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and anti-mutator, *Cell* 90 (1997) 43–53.
- [53] S. Altuvia, A. Zhang, L. Argaman, A. Tiwari, G. Storz, The *Escherichia coli* OxyS regulatory RNA represses *hflA* translation by blocking ribosome binding, *EMBO J.* 17 (1998) 6069–6075.
- [54] L. Argaman, S. Altuvia, *hflA* repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense–target RNA complex, *J. Mol. Biol.* 300 (2000) 1101–1112.
- [55] A. Zhang, K.M. Wassarman, J. Ortega, A.C. Steven, G. Storz, The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs, *Mol. Cell* 9 (2002) 11–22.
- [56] E.G. Wagner, K. Flardh, Antisense RNAs everywhere? *Trends Genet.* 18 (2002) 223–226.
- [57] M. Givskov, S. Molin, Copy mutants of plasmid R1: effects of base pair substitutions in the *copA* gene on the replication control system, *Mol. Gen. Genet.* 194 (1984) 286–292.
- [58] C.L. Beisel, T.B. Updegrove, B.J. Janson, G. Storz, Multiple factors dictate target selection by Hfq-binding small RNAs, *EMBO J.* 31 (2012) 1961–1974.
- [59] Q.J. Li, J. Chau, P.J. Ebert, G. Sylvester, H. Min, G. Liu, R. Braich, M. Manoharan, J. Soutschek, P. Skare, L.O. Klein, M.M. Davis, C.Z. Chen, miR-181a is an intrinsic modulator of T cell sensitivity and selection, *Cell* 129 (2007) 147–161.
- [60] S. Guil, J.F. Caceres, The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a, *Nat. Struct. Mol. Biol.* 14 (2007) 591–596.
- [61] M.A. Newman, J.M. Thomson, S.M. Hammond, Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing, *RNA* 14 (2008) 1539–1549.
- [62] I. Heo, C. Joo, J. Cho, M. Ha, J. Han, V.N. Kim, Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA, *Mol. Cell* 32 (2008) 276–284.
- [63] S.R. Viswanathan, G.Q. Daley, R.I. Gregory, Selective blockade of microRNA processing by Lin28, *Science* 320 (2008) 97–100.
- [64] M. Trabucchi, P. Briata, M. Garcia-Mayoral, A.D. Haase, W. Filipowicz, A. Ramos, R. Gherzi, M.G. Rosenfeld, The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs, *Nature* 459 (2009) 1010–1014.
- [65] D. Glatfield, G. Le Martelot, C.E. Vejnar, D. Gerlach, O. Schaad, F. Fleury-Olela, A.L. Ruskeepaa, M. Oresic, C.C. Esau, E.M. Zdobnov, U. Schibler, Integration of microRNA miR-122 in hepatic circadian gene expression, *Genes Dev.* 23 (2009) 1313–1326.
- [66] R. Fragoso, T. Mao, S. Wang, S. Schaffert, X. Gong, S. Yue, R. Luong, H. Min, Y. Yashiro-Ohtani, M.M. Davis, W. Pear, C.Z. Chen, Modulating the strength and threshold of NOTCH oncogenic signals by mir-181a-1/b-1, *PLoS Genet.* 8 (8) (2012) e1002855.
- [67] R.M. O'Connell, A. Chaudhuri, D.S. Rao, W.S. Gibson, A.B. Balazs, D. Baltimore, MicroRNAs enriched in hematopoietic stem cells differentially regulate long-term hematopoietic output, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 14235–14240.
- [68] Y.W. Kong, I.G. Cannell, C.H. de Moor, K. Hill, P.G. Garside, T.L. Hamilton, H.A. Meijer, H.C. Dobbyn, M. Stoneley, K.A. Spriggs, A.E. Willis, M. Bushell, The mechanism of micro-RNA-mediated translation repression is determined by the promoter of the target gene, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 8866–8871.
- [69] C.M. Hammell, I. Lubin, P.R. Boag, T.K. Blackwell, V. Ambros, nhl-2 modulates microRNA activity in *Caenorhabditis elegans*, *Cell* 136 (2009) 926–938.
- [70] M.M. Janas, B. Wang, A.S. Harris, M. Aguiar, J.M. Shaffer, Y.V. Subrahmanyam, M.A. Behlke, K.W. Wucherpfennig, S.P. Gygi, E. Gagnon, C.D. Novina, Alternative RISC assembly: binding and repression of microRNA–mRNA duplexes by human Ago proteins, *Rna* 18 (2012) 2041–2055.
- [71] G.S. Tan, B.G. Garchow, X. Liu, J. Yeung, J.P.t. Morris, T.L. Cuellar, M.T. McManus, M. Kiriakidou, Expanded RNA-binding activities of mammalian Argonaute 2, *Nucleic Acids Res.* 37 (2009) 7533–7545.
- [72] D.G. Zisoulis, Z.S. Kai, R.K. Chang, A.E. Pasquinelli, Autoregulation of microRNA biogenesis by let-7 and Argonaute, *Nature* 486 (2012) 541–544.
- [73] S. Kuchen, W. Resch, A. Yamane, N. Kuo, Z. Li, T. Chakraborty, L. Wei, A. Laurence, T. Yasuda, S. Peng, J. Hu-Li, K. Lu, W. Dubois, Y. Kitamura, N. Charles, H.W. Sun, S. Muljo, P.L. Schwartzberg, W.E. Paul, J. O'Shea, K. Rajewsky, R. Casellas, Regulation of microRNA expression and abundance during lymphopoiesis, *Immunity* 32 (2010) 828–839.
- [74] J. Krol, I. Loedige, W. Filipowicz, The widespread regulation of microRNA biogenesis, function and decay, *Nat. Rev. Genet.* 11 (2010) 597–610.
- [75] S. Chatterjee, H. Grosshans, Active turnover modulates mature microRNA activity in *Caenorhabditis elegans*, *Nature* 461 (2009) 546–549.
- [76] X. Cai, C.H. Hagedorn, B.R. Cullen, Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs, *RNA* 10 (2004) 1957–1966.
- [77] S.W. Chi, J.B. Zang, A. Mele, R.B. Darnell, Argonaute HITS-CLIP decodes microRNA–mRNA interaction maps, *Nature* 460 (2009) 479–486.